

## Contenuti

- Architettura di Internet
- Principi di interconnessione e trasmissione

**Tecnologie delle reti di calcolatori**

- World Wide Web
- Posta elettronica
- ➔ **Motori di ricerca**

**Servizi Internet**  
*(come funzionano e come usarli)*

- Antivirus
- Personal firewall

**Servizi Internet**  
*(come difendersi)*

## Modulo 1: Motori di ricerca

## **Motori di ricerca**

- Un'applicazione che aiuta nella ricerca di documenti su Web
- La ricerca è basata su l'uso di parole chiave (keyword)
- Il risultato è una lista di risorse Web che contengono la/e parola/e chiave richiesta/e

- Il più famoso è:

[www.google.it](http://www.google.it)

## **Interfaccia dei motori di ricerca**

- Ogni motore di ricerca contiene una zona scrivibile nella quale digitare la richiesta (query)
- Dopo di che premendo il tasto SEARCH (o tasti dai nomi analoghi) il motore avvia la ricerca dei documenti presenti sul Web che si adeguano alla richiesta

## Interfaccia dei motori di ricerca

- La query più semplice è costituita da un semplice termine da cercare.
- Premuto il tasto Search si ottiene in risposta la lista di tutti i documenti che contengono il termine indicato.
- Ogni risposta comprende:
  - il titolo di una pagina WWW
  - l'indirizzo (cliccabile) della pagina
  - un estratto della pagina, per avere un'idea del contenuto
  - eventuali altre informazioni, come p.es. la dimensione in byte della pagina, la data di creazione o ultimo aggiornamento, ecc.

## Esempi di siti

- <http://www.google.com>
- <http://www.altavista.com>
- <http://www.excite.com>
- <http://www.yahoo.com>

## Strumenti dei motori di ricerca

- Solitamente un motore di ricerca utilizza una serie di programmi *spider* che navigano continuamente ed indipendentemente tra i siti Web e ritornano quanti più documenti possibile
- Un altro programma, chiamato *indexer*, legge i documenti e crea un indice basato sulle parole contenute in ogni documento
- Ogni motore di ricerca utilizza un programma proprietario per indicizzare i risultati in modo opportuno

## Aspettative utente

- Ci si aspetta che i documenti più interessanti o maggiormente collegati alla lista di parole chiave fornita si trovino in cima alla lista
- Questo può non succedere se:
  - Il programma di indicizzazione non è efficiente
  - La selezione di parole chiave fornita dall'utente è ambigua o troppo generica

## Motori di ricerca

- Se la parola chiave è troppo generica (è associata a più significati diversi) o se la richiesta dell'utente è ambigua, si può avere un sovraccarico di informazioni:
  - Il motore di ricerca ritorna una lista troppo lunga di risultati
- Una soluzione è quella di raffinare la ricerca, aggiungendo parole chiave o utilizzando gli operatori booleani forniti dalle opzioni di ricerca avanzata del motore.

## Operatori booleani

- Il matematico inglese George Boole (1815-1864) fondò un campo della matematica e della filosofia chiamato logica simbolica
- Il suo nome è rimasto legato ad un insieme di operatori che sono molto utili e molto presenti nel campo dell'informatica e che si chiamano operatori booleani
- Nell'ambito dei motori di ricerca gli operatori booleani sono utili per definire operazioni di ricerca avanzate

## Operatori booleani (*cont.*)

- Gli operatori booleani di base sono:
  - **AND**
  - **OR**
  - **NOT**
- Essi vengono applicati a uno (nel caso del NOT) o due (nel caso di AND e OR) argomenti e ritornano dei valori di verità (VERO o FALSO)

## Come scrivere una query

- La modalità di ricerca all'interno del Web varia a seconda del motore di ricerca usato.
  - Solitamente, le parole di una query vengono cercate all'interno dei documenti in qualunque ordine e non necessariamente tutte insieme
  - Inoltre esse vengono filtrate per eliminare parole poco significative (articoli), punteggiatura, ... Questa operazione è detta **stoplist removal**

Es. sto cercando: *La divina commedia; opera di Dante*

Dopo l'operazione di stoplist removal ottengo: *divina commedia opera Dante*

## Come scrivere una query

- Esse vengono anche filtrate, tramite opportuni algoritmi, per unificare le parole con radice comune (e.g. Canto, cantare, cantante, cantano, canti...). Questa operazione è detta **stemming**
- Ad esempio se un utente cercherà la parola "Danza" il motore ricercherà e proporrà fra i risultati anche le parole "Danzatrice" , "Danzatore" etc...

## Come scrivere una query

- Gli indirizzi che vengono riportati corrispondono quindi a pagine che contengono:
  - Tutte o alcune delle parole della query
  - Parole simili a quelle della query (p.es. un'altra voce dello stesso verbo)
  - A volte, parole con significato correlato a quello delle parole della query
- Gli indirizzi vengono inoltre ordinati per **rilevanza** (i criteri di rilevanza variano da motore di ricerca a motore di ricerca)

## Come scrivere una query

- Meglio una query con molte parole che una query meno specifica
- E' possibile anche inserire come query una domanda vera e propria. (grazie allo stop word removal)
- Termini obbligatori:
  - In alcuni casi io voglio essere sicuro che le parole che scelgo siano comprese nei testi trovati (tipo l'AND logico)
  - Precedo ogni parola obbligatoria da un segno +

## Come scrivere una query

- Esclusione di termini:
  - Talvolta una stessa parola ha due significati diversi. Se vogliamo essere sicuri di ottenere solo risposte relative ad un significato possiamo escludere esplicitamente parole che ci ricondurrebbero all'altro
  - Esempio: se cerco informazioni relative al calcio (nel senso di elemento chimico) posso provare ad escludere i risultati relativi allo sport nel seguente modo:
    - Calcio –sport –soccer –pallone
  - Anche se forse una ricerca del tipo
    - Calcio chimica
    - Potrebbe essere più efficace



## Come scrivere una query

- Ricerca di frasi o di parole adiacenti:
  - Racchiudo la frase tra virgolette
  - Mi verranno restituiti solo i documenti in cui la frase compare in modo esatto (con le parole nell'ordine in cui le ho scritte).